

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-052984

(43)Date of publication of application : 26.02.1999

(51)Int.Cl. G10L 3/00
G10L 3/00
G06F 3/16

(21)Application number : 09-215308

(71)Applicant : FUJITSU LTD

(22)Date of filing : 08.08.1997

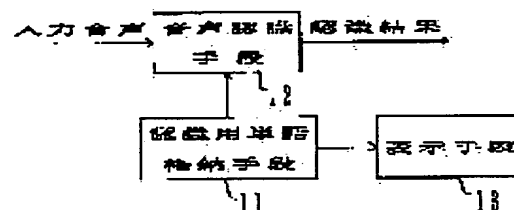
(72)Inventor : KIMURA AKIHIRO
SATO YASUO
KOIDE TOMIO

(54) RECOGNITION OBJECT DISPLAY SYSTEM IN SPEECH RECOGNITION DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To eliminate the need for mouse operation and to improve an effect of a speech recognition device by automatically display words stored in a word storage means for recognition for helping a speech input to the speech recognition device by user.

SOLUTION: A word storage means 11 for recognition stores all words as recognition object of an input speech. Based on the feature of the input speech, a speech recognition means 12 determines a most similar word to the input speech from the words for recognition, and outputs it as a recognition result. Further, a display means 13 automatically displays the words stored in the word storage means 11 for recognition to help the speech input to the speech input system by the user. For example, when a word is found which should be inputted in the words displayed on a display screen, the word is speech- inputted. The input speech is collated with all the recognition words stored in the word storage means 11 for recognition based on the feature of the speech, a most similar word to the input speech is outputted from the speech recognition means 12 as a result of the recognition.



LEGAL STATUS

[Date of request for examination]

01.07.2004

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision
of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-52984

(43) 公開日 平成11年(1999) 2月26日

(51) Int. Cl. ⁶	識別記号	F I
G 1 0 L 3/00	5 7 1	G 1 0 L 3/00 5 7 1 H
		S
G 0 6 F 3/16	3 2 0	G 0 6 F 3/16 3 2 0 H

審査請求 未請求 請求項の数16 O L (全 18 頁)

(21) 出願番号 特願平9-215308

(22) 出願日 平成9年(1997) 8月8日

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72) 発明者 木村 晋太

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(72) 発明者 佐藤 泰雄

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(72) 発明者 小出 富夫

東京都立川市曙町1-21-11 クリエイト
システム開発株式会社内

(74) 代理人 弁理士 大菅 義之 (外1名)

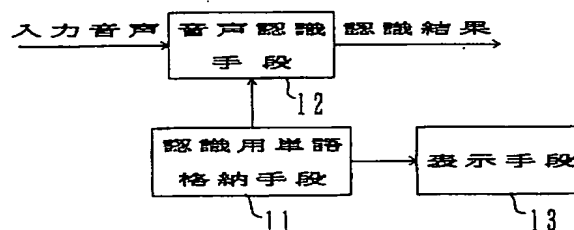
(54) 【発明の名称】 音声認識装置における認識対象表示方式

(57) 【要約】

【課題】 ユーザが認識対象単語をよく知らない場合においても、音声認識装置の操作性を向上させ、その効用を大きくする。

【解決手段】 認識対象単語を格納する手段11と、入力される音声の特徴に基づいて、手段11に格納されている全ての単語の中で入力音声に最も類似する単語を求めて認識結果とする手段12と、ユーザの音声入力を助けるために、手段11に格納されている単語を自動的に表示する手段13とを備える。

第1の発明の原理構成ブロック図



【特許請求の範囲】

【請求項 1】 入力される音声認識する音声認識装置において、

該入力音声の認識対象としての単語を格納した認識用単語格納手段と、

入力される音声の特徴に基づいて、該認識用単語格納手段に格納されている全ての認識対象単語の中から入力音声に最も類似する単語を求めて認識結果とする音声認識手段と、

前記音声認識装置のユーザの音声入力を助けるために、前記認識用単語格納手段に格納されている単語を自動的に表示する表示手段とを備えることを特徴とする音声認識装置における認識対象表示方式。

【請求項 2】 前記表示手段が、前記認識用単語格納手段に格納されている認識対象単語をディスプレイ画面上で自動的にスクロールさせながら表示することを特徴とする請求項 1 記載の音声認識装置における認識対象表示方式。

【請求項 3】 前記表示手段が、前記認識用単語格納手段に格納されている認識対象単語をディスプレイ画面上で自動的に改ページさせながら表示することを特徴とする請求項 1 記載の音声認識装置における認識対象表示方式。

【請求項 4】 前記表示手段が、前記認識用単語格納手段に格納されている認識対象単語を、少なくとも 1 単語ずつの単位で順次読み出して、ディスプレイ画面上に自動的に表示することを特徴とする請求項 1 記載の音声認識装置における認識対象表示方式。

【請求項 5】 前記表示手段が、前記スクロール、または改ページのタイミング、あるいは少なくとも 1 単語ずつを単位とする順次読み出しのタイミングをクロック信号に基づいて決定することを特徴とする請求項 2、3、または 4 記載の音声認識装置における認識対象表示方式。

【請求項 6】 前記表示手段が、前記認識用単語格納手段に格納されている認識対象単語を、少なくとも 1 単語以上ずつの単位で単語に付けられた番号に対応する乱数の値に基づいて読み出して、ディスプレイ画面上に自動的に表示することを特徴とする請求項 1 記載の音声認識装置における認識対象表示方式。

【請求項 7】 入力される音声認識する音声認識装置において、

入力される音声の特徴に基づいて、現在の入力音声の認識対象となる単語が格納される認識用単語格納手段に格納されている全ての認識対象単語の中から入力音声に最も類似した単語を求めて認識結果とする音声認識手段と、

前記音声認識装置のユーザの音声入力を助けるために、入力音声の認識対象となる可能性のある単語を格納した単語格納手段に格納されている単語を自動的に表示する

表示手段と、

該表示手段によって現在までに表示された単語の中で最近表示された単語を、前記認識対象となる単語として抽出し、前記認識用単語格納手段に格納する表示単語抽出手段とを備えることを特徴とする音声認識装置における認識対象表示方式。

【請求項 8】 前記表示手段が、前記認識用単語格納手段に格納されている認識対象単語をディスプレイ画面上で自動的にスクロールさせながら表示することを特徴とする請求項 7 記載の音声認識装置における認識対象表示方式。

【請求項 9】 前記表示手段が、前記認識用単語格納手段に格納されている認識対象単語をディスプレイ画面上で自動的に改ページさせながら表示することを特徴とする請求項 7 記載の音声認識装置における認識対象表示方式。

【請求項 10】 前記表示手段が、前記認識用単語格納手段に格納されている認識対象単語を、少なくとも 1 単語ずつの単位で順次読み出して、ディスプレイ画面上に自動的に表示することを特徴とする請求項 7 記載の音声認識装置における認識対象表示方式。

【請求項 11】 前記表示手段が、前記スクロール、または改ページのタイミング、あるいは少なくとも 1 単語ずつを単位とする順次読み出しのタイミングをクロック信号に基づいて決定することを特徴とする請求項 8、9、または 10 記載の音声認識装置における認識対象表示方式。

【請求項 12】 前記表示手段が、前記認識用単語格納手段に格納されている認識対象単語を、少なくとも 1 単語以上ずつの単位で単語に付けられた番号に対応する乱数の値に基づいて読み出して、ディスプレイ画面上に自動的に表示することを特徴とする請求項 7 記載の音声認識装置における認識対象表示方式。

【請求項 13】 入力される音声の特徴に基づいて、認識用単語格納手段に格納されている全ての認識対象単語の中から入力音声に最も類似する単語を求めて認識結果とする第 1 の手順と、

ユーザの音声入力を助けるために、前記認識用単語格納手段に格納されている単語を自動的に表示する第 2 の手順とをコンピュータに実行させるプログラムを記憶したコンピュータ読み取り可能な記憶媒体。

【請求項 14】 入力される音声の特徴に基づいて、現在の入力音声の認識対象となる単語が格納される認識用単語格納手段に格納されている全ての認識対象単語の中から入力音声に最も類似した単語を求めて認識結果とする第 1 の手順と、

前記音声認識装置のユーザの音声入力を助けるために、入力音声の認識対象となる可能性のある単語を格納した単語格納手段に格納されている単語を自動的に表示する第 2 の手順と、

該第2の手順によって現在までに表示された単語の中で最近表示された単語を、前記認識対象となる単語として抽出し、前記認識用単語格納手段に格納する第3の手順とをコンピュータに実行させるプログラムを記憶したコンピュータ読み取り可能な記憶媒体。

【請求項15】 入力される音声認識する音声認識装置において、

入力音声の認識対象としての単語を格納する第1の手順と、

入力される音声の特徴に基づいて、格納されている全ての認識対象単語の中から入力音声に最も類似する単語を求めて認識結果とする第2の手順と、

ユーザの音声入力を助けるために、格納されている単語を自動的に表示する第3の手順とからなることを特徴とする音声認識装置における認識対象表示方法。

【請求項16】 入力される音声の特徴に基づいて、現在の入力音声の認識対象となる全ての認識対象単語の中から入力音声に最も類似した単語を求めて認識結果とする第1の手順と、

前記音声認識装置のユーザの音声入力を助けるために、入力音声の認識対象となる可能性のある単語を格納した単語格納手段に格納されている単語を自動的に表示する第2の手順と、

該第2の手順によって現在までに表示された単語の中で最近表示された単語を、前記認識対象となる単語として抽出し、前記認識用単語格納手段に格納する第3の手順とを備えることを特徴とする音声認識方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は音声認識装置に係り、更に詳しくは音声認識の認識対象となる単語をディスプレイ画面上に自動的に表示することによってユーザの音声入力を助ける、音声認識装置における認識対象表示方式に関する。

【0002】

【従来の技術】図22は音声認識装置の従来例の構成ブロック図である。同図において音声認識装置は音声が入力され、電気信号に変換されるマイク1、例えば高速フーリエ変換や、線形予測分析などを用いて入力された音声进行分析する分析部2、音声の分析結果から特徴を抽出する特徴抽出部3、認識対象としての単語を格納する単語辞書部4、単語辞書部4に格納されている全ての単語と入力音声を、特徴抽出部3によって抽出された特徴に基づいて照合し、入力音声に最も類似した単語を求める単語照合部5、単語照合部5によって出力される単語の認識結果と、単語辞書部4に格納されている単語を表示する表示部6、単語辞書部4に格納されている単語の表示部6による表示に際して、単語辞書部4に格納されている単語の数が多く、1つの画面上に全ての単語を表示することができないため、表示単語のスクロールを行う

スクロール制御部7、スクロール制御部7にスクロールの指示を与えるためのマウス8を備えている。

【0003】単語照合部5では、入力音声と単語辞書部4に格納されている全ての単語の照合が行われるが、その手法としては、例えば入力音声の時間方向の伸び縮みを吸収するための手法としてのDPマッチング法、または単語音声の確率モデルを使用する隠れマルコフモデル法(HMM法)などが用いられる。この照合の結果としての認識結果は、前述のように表示部6に表示されるが、本発明が対象とする音声認識装置では、認識結果と共に音声認識の対象としての単語辞書部4に格納されている単語が、同時に表示部6によって表示される。

【0004】本発明が対象とする音声認識装置では、認識対象としての単語を音声認識装置のユーザが必ずしもよく知らないということを前提としている。音声認識装置が、例えばコマンドの入力用に使用されるものである時、ユーザはコマンドをよく知っていないため、表示部6によって表示されるコマンドを見ながら、入力すべきコマンドが表示された時にそのコマンドを音声入力することによって、コマンドの入力を実行できるものとする。そのためにスクロール制御部7によって、単語辞書部4に格納されている全てのコマンドをディスプレイ画面上でスクロールさせながら表示して、ユーザのコマンド入力を助けることになる。

【0005】

【発明が解決しようとする課題】前述のように、本発明が対象とする音声認識装置では、どのような単語が認識対象となっているかをユーザはあまりよく知らず、図22においてはマウス8によってディスプレイ画面上に表示されている認識対象単語、例えばコマンドをスクロールさせながら入力すべきコマンドを探して、音声入力を行うことが必要であった。このように従来例ではマウスを使って入力すべき単語を探しているためディスプレイ画面上にその単語を見つけた時にマウスによってその単語をクリックすれば、その単語、例えばコマンドの入力を行うことができることになり、そもそも音声認識を用いる必要性がはっきりしないという問題点があった。

【0006】本発明は、音声認識の対象となる単語を、例えばディスプレイ画面上で自動的にスクロールさせながら表示することにより、マウス操作を行うことを不必要にさせ、音声認識装置の効用を大きくすることである。

【0007】

【課題を解決するための手段】図1および図2は本発明の原理構成ブロック図である。これらの図は入力される音声認識する音声認識装置、特に認識対象の単語をディスプレイ画面上に表示する音声認識装置の原理構成ブロック図である。

【0008】図1は第1の発明の原理構成ブロック図である。同図において、認識用単語格納手段11は入力音

声の認識対象としての単語、例えばコマンドの全てを格納しているものである。音声認識手段 12 は入力される音声の特徴に基づいて、認識用単語格納手段 11 に格納されている全ての認識対象単語、例えばコマンドの中から最も入力音声に類似した単語を求めて、認識結果として出力するものである。更に表示手段 13 は音声認識装置のユーザの音声入力を助けるために、認識用単語格納手段 11 に格納されている単語を自動的に表示するものである。

【0009】第 1 の発明においては表示手段 13 によって、例えばディスプレイ画面上で認識対象単語が自動的に表示され、ユーザは表示されている単語の中で自分が入力すべき単語を見つけた時点でその単語の音声入力を行う。入力された音声は、その特徴に基づいて認識用単語格納手段 11 に格納されている全ての認識対象単語と照合され、最も入力音声に類似した単語が認識結果として音声認識手段 12 から出力される。

【0010】第 1 の発明の実施の形態においては、表示手段 13 による単語の自動的表示に様々な方式が用いられる。その方式としては、例えば認識用単語格納手段 11 に格納されている全ての単語が、自動的にスクロール制御によって表示されたり、自動的に改ページされて表示されたり、あるいは少なくとも 1 単語ずつの単位で順次に読み出されて表示されたり、単語に付けられた番号に対応する乱数の値に基づいて読み出されて表示されたりすることができる。これによってユーザがマウス操作を行う必要はなくなる。

【0011】図 2 は第 2 の発明の原理構成ブロック図である。同図において単語格納手段 15 は、入力音声の認識対象となる可能性のある単語を全て格納しているものである。認識用単語格納手段 16 は、現在の時点の入力音声の認識対象となる単語を格納するものである。音声認識手段 17 は、入力される音声の特徴に基づいて、認識用単語格納手段 16 に格納されている全ての単語の中から入力音声に最も類似した単語を求めて、認識結果として出力するものである。

【0012】表示手段 18 は音声認識装置のユーザの音声入力を助けるためのものであり、単語格納手段 15 に格納されている単語、すなわち認識対象となる可能性のある単語を自動的に表示するものである。表示単語抽出手段 19 は、表示手段 18 によって現在までに表示された単語の中で、最近表示された単語を認識対象となる単語として抽出し、認識用単語格納手段 16 に格納するものである。

【0013】第 2 の発明においては、第 1 の発明におけると同様に単語格納手段 15 に格納されている全ての単語が例えばスクロール制御や、改ページ制御、少なくとも 1 単語ずつの単位による順次読み出しの制御によってディスプレイ画面上に自動的に表示される。表示手段 18 によって表示された単語の中で最近表示された単語が

表示単語抽出手段 19 によって抽出され、認識用単語格納手段 16 に格納されて、現在の入力音声の認識対象となる単語として使用される。入力音声は、この認識用単語格納手段 16 に格納されている単語と照合されて、入力音声と最も類似した単語が認識結果として音声認識手段 17 によって出力されることになる。これによって第 2 の発明では認識対象は、最近表示された単語に絞られることになる。ユーザが入力する単語の種類が比較的少なく、かつ最近表示された単語と同じ単語が多い場合には、このように認識対象を限定することによって音声認識の性能を向上させることができる。

【0014】また、本発明は入力される音声の特徴に基づいて、認識用単語格納手段に格納されている全ての認識対象単語の中から入力音声に最も類似する単語を求めて認識結果とする第 1 の手順と、ユーザの音声入力を助けるために、前記認識用単語格納手段に格納されている単語を自動的に表示する第 2 の手順とをコンピュータに実行させるプログラムを記憶したコンピュータ読み取り可能な記憶媒体を構成する。

【0015】さらに、本発明は入力される音声の特徴に基づいて、現在の入力音声の認識対象となる単語が格納される認識用単語格納手段に格納されている全ての認識対象単語の中から入力音声に最も類似した単語を求めて認識結果とする第 1 の手順と、前記音声認識装置のユーザの音声入力を助けるために、入力音声の認識対象となる可能性のある単語を格納した単語格納手段に格納されている単語を自動的に表示する第 2 の手順と、該第 2 の手順によって現在までに表示された単語の中で最近表示された単語を、前記認識対象となる単語として抽出し、前記認識用単語格納手段に格納する第 3 の手順とをコンピュータに実行させるプログラムを記憶したコンピュータ読み取り可能な記憶媒体を構成する。

【0016】

【発明の実施の形態】図 3 は本発明における音声認識装置の第 1 の実施形態の構成ブロック図である。同図において従来例の図 2 と同じ動作をする部分には同じ符号を付してある。図 2 と異なる部分はスクロール制御部 21 と、クロック部 22 である。

【0017】この第 1 の実施形態においては、単語辞書部 4 に格納されている認識対象単語が表示部 6 によって表示されるが、その表示はスクロール制御部 21 によって制御される。このスクロール動作は、図 2 と異なって自動的に行われる。すなわちクロック部 22 の発生するクロック信号によって、例えば 1 クロック毎にスクロールが行われ、単語辞書部 4 に格納されている認識対象単語が m 単語、常に画面上に表示される形式でスクロール制御される。

【0018】図 4 は図 3 の第 1 の実施形態における全体処理のフローチャート、図 5 は図 4 における認識プロセスの処理フローチャートである。図 4 において処理が開

始されると、まずステップS1で認識プロセスが起動された後に、ステップS2で単語辞書部4の内容のスクロール表示が開始され、ステップS3でディスプレイ画面の、例えばm単語のスクロール表示の1行スクロールが繰り返される。

【0019】図5の認識プロセスでは、まずステップS5でマイク1からの音声入力を受け取られ、ステップS6でその音声分析部2によって分析され、ステップS7でその分析結果から特徴抽出部3によって入力音声の特徴が抽出され、ステップS8で、その特徴が単語照合部5によって、単語辞書部4に格納されている全ての単語と照合され、ステップS9でその照合結果、すなわち最も類似した単語が認識結果としてアプリケーションに通知されて、マイク1からの次の音声入力を待つ状態となる。

【0020】図6は図3の第1の実施形態における複数単語の表示イメージの説明図である。同図においては山手線の駅名が連続した順序で1つの画面上に5つ表示され、クロック毎に1駅ずつ上にずれる形式でスクロール表示される。

【0021】図7は図3におけるスクロール制御部21の詳細構成ブロック図である。同図においてスクロール制御部21は、1つの画面上での表示開始位置、例えば一番上の位置の単語の番号を格納する表示開始単語番号レジスタ25、表示開始単語番号レジスタ25の出力Aと1つの画面上に表示される表示単語数から1を減算したBとを加算する加算器26、加算器26の出力を1つの画面上での表示終了位置、例えば一番下の単語の番号として格納する表示終了単語番号レジスタ27、図3の単語辞書部4に格納されている全ての単語数Aと加算器26の出力Bとを比較して、BがAを越えた時にカウンタ29をリセットする比較器28、クロック信号によってカウント値がインクリメントされ、表示開始単語番号レジスタ25にカウント値を出力するカウンタ29から構成されている。

【0022】図8はクロック信号に対応する表示開始単語番号と表示終了単語番号の説明図である。同図において1つの画面上に表示される単語の数はmであり、図3の単語辞書部4に格納されている単語の数はnであるものと仮定している。クロックが進む毎に表示開始単語番号が大きくなり、n番目の単語が表示開始位置にきた時には表示終了単語番号は $n+m-1$ となる。そして次のクロックでは表示開始単語番号は1となり、クロックの進行と共に単語表示のスクロール制御が繰り返される。

【0023】図9は音声認識装置の第2の実施形態の構成ブロック図である。同図を図3の第1の実施形態と比較すると、スクロール制御部21に代わって一単語読み出し部31が備えられている点だけが異なっている。この第2の実施形態においては、単語辞書部4に格納されている音声認識の対象語の表示にあたってスクロール制

御が用いられず、単語辞書部4の内容が1単語ずつ一単語読み出し部31に従って読み出され、表示部6によって順次表示される。この1単語ずつの読み出しは、クロック部22が出力するクロック信号に同期して行われる。なおここでは1単語ずつを読み出すものとしたが、実施形態としては1単語ずつに限定されず、例えば2単語ずつとか、3単語ずつのように語数を増やしたり、例えばユーザからの指示に応じて読み出すべき単語数を変化させたりすることも当然可能である。

【0024】図10は第2の実施形態における全体処理のフローチャートである。同図において処理が開始されると、まずステップS11で認識プロセスが起動される。この認識プロセスの処理は図5における同じであるので、その説明を省略する。

【0025】続いてステップS12でポインタが単語辞書部4の先頭の単語を指すように置かれ、ステップS13でポインタで指定された単語が表示され、ステップS14でポインタが1単語分進められ、ステップS15でポインタが単語辞書部4に格納されている単語の最後に達したか否かが判定され、まだ達していない場合にはステップS13以降の処理が繰り返される。単語辞書部4に格納されている単語の最後に達したと判定されると、ステップS12でポインタが単語辞書の先頭に戻された後に、ステップS13以降の処理が繰り返される。

【0026】図11は第2の実施形態において認識対象語が1単語ずつ順次読み出される場合の表示イメージの説明図である。同図に示すように、山手線の駅名がクロック毎に1つずつ読み出されて表示される。

【0027】図12は図9の一単語読み出し部31の構成ブロック図である。同図において一単語読み出し部31は読み出すべき単語番号を指定するカウンタ33、単語辞書部4から読み出された単語が一時的に格納され、その単語が表示部6に出力される単語バッファ34から構成されている。カウンタ33には、カウント値をインクリメントさせるためのクロック信号、および単語辞書部4に格納されている単語数nが入力され、カウンタ33のカウント値はクロック信号に同期してインクリメントされ、nに達した時点で自動的にリセットされる。

【0028】図13は音声認識装置の第3の実施形態の構成ブロック図である。同図を図9の第2の実施形態と比較すると、クロック部22に代わって乱数発生部36が備えられている点が異なっている。一単語読み出し部35は、乱数発生部36によって発生される乱数の値の単語番号を持つ単語を単語辞書部4から一単語ずつ読み出して、その単語が表示部6によって表示される。

【0029】図14は図13の第3の実施形態における全体処理のフローチャートである。同図において処理が開始されると、ステップS16で認識プロセスが起動される。この認識プロセスによる処理は図5と同じであるので、その説明を省略する。続いてステップS17で単

語辞書部 4 から、乱数発生部 36 が出力する乱数に基づいて、認識対象単語が一単語ランダムに取り出され、ステップ S 18 でその単語が表示され、以後ステップ S 17 および 18 の処理が繰り返される。

【0030】図 15 は図 13 における一単語読み出し部 35 の構成ブロック図である。同図において一単語読み出し部 35 は、読み出すべき単語番号を単語辞書部 4 に与える単語番号バッファ 38、単語辞書部 4 から読み出された単語を一時的に格納し、表示部 6 に出力するための単語バッファ 39 から構成されている。そして乱数発生部 36 によって出力される乱数が単語番号バッファ 38 に格納され、その乱数によって指定される単語番号を持つ認識対象単語が単語辞書部 4 から読み出されて、単語バッファ 39 を介して表示部 6 によって表示されることになる。

【0031】図 16 は音声認識装置の第 4 の実施形態の構成ブロック図である。同図を第 1 の実施形態を示す図 3 と比較すると、図 3 の構成要素に加えて認識用一時単語辞書部 41 と表示単語抽出部 42 が追加され、単語辞書部 4 の格納内容はスクロール制御部 21 のみに与えられる点が異なっている。この第 4 の実施形態では、単語照合部 5 において用いられる認識対象の単語は単語辞書部 4 に格納されている単語ではなく、認識用一時単語辞書部 41 に格納されている単語だけに限定される。この認識用一時単語辞書部 41 に格納される認識用単語は、表示単語抽出部 42 によって抽出された単語である。

【0032】この表示単語抽出部 42 の動作については更に後述するが、抽出部 42 は表示部 6 によって最近表示された単語を抽出して、認識用単語として認識用一時単語辞書部 41 に格納する。表示部 6 によって表示される単語は、単語照合部 5 の出力、すなわち認識結果としての単語と、スクロール制御部 21 によってスクロール制御される、単語辞書部 4 に格納されている単語のうちの一部であるが、表示単語抽出部 42 によって抽出される認識用単語には単語照合部 5 の出力、すなわち認識結果としての単語は含まれない。

【0033】本発明が対象とする音声認識装置は、前述のようにユーザが音声認識の対象語をよく知らないような場合を対象としている。例えばコマンドを入力するための音声認識装置では、ユーザはコマンドの名前をあまり知らず、表示部 6 によって表示されるコマンドの中から自分の入力すべきコマンドを知って、それを音声入力することになる。

【0034】このような場合にはユーザが入力したいコマンドの種類はそれほど多くなく、表示部 6 によって最近表示されたものを再度入力することが多くなるため、認識用一時単語辞書部 41 に格納され、単語照合部 5 によって音声認識の対象として用いられる単語を最近表示されたものに限定することによって認識の速度を向上させ、また認識エラーを減少させることができる。

【0035】一般に音声認識においては入力される音声 が 1 つの音毎に正しく認識されることが要求される。そこで認識の対象となる単語が多ければ多いほど、単語照合部 5 による照合に時間がかかり、またマイクの特性や人の声のくせなどのために、正しい認識を行うことが困難になり、認識エラーが生ずる可能性がある。この第 4 の実施形態においては、単語照合部 5 における認識対象を表示部 6 によって最近表示されたものに限定することによって、認識に要する時間を減少させ、かつ認識エラーを少なくすることが可能となる。

【0036】図 17 は第 4 の実施形態における全体処理のフローチャート、図 18 は認識プロセスによる処理のフローチャートである。図 17 を第 1 の実施形態に対する図 4 と比較すると、ステップ S 3 において 1 行のスクロールが行われた後に、ステップ S 20 で最近表示された単語を抽出し、認識用一時単語辞書部 41 に格納する処理が追加されている点が異なっている。

【0037】図 18 を第 1 ～第 3 の実施形態に対する図 5 と比較すると、ステップ S 8 で入力音声の特徴が単語辞書部 4 に格納されている単語と照合される代わりに、ステップ S 21 で認識用一時単語辞書部 41 に格納されている単語と照合される点が異なっている。

【0038】図 19 は図 16 における表示単語抽出部 42 の構成ブロック図である。図 20 はその中でのリングバッファにおける処理内容の説明図である。図 19 において表示単語抽出部 42 は、表示部 6 に表示されている単語を格納する表示単語バッファ 44、表示単語バッファ 44 の出力を格納するリングバッファ 45、リングバッファ 45 の内容を読み出し、認識用一時単語辞書部 41 に出力する読み出し部 46 から構成されている。

【0039】図 20 においてリングバッファ 45 のバッファメモリには、例として山手線の駅の名前が 7 つ格納されている。表示部 6 に現在表示され、図 19 の表示単語バッファ 44 に格納された単語は、バッファメモリの内部で一番古いデータに上書きする形式で書込まれる。そしてバッファメモリに格納されている単語、ここでは 7 つの駅の名前が、読み出し部 46 を介して認識用一時単語辞書部 41 に与えられることになる。

【0040】図 21 は音声認識装置の第 5 の実施形態の構成ブロック図である。同図を第 4 の実施形態を示す図 16 と比較すると、スクロール制御部 21、クロック部 22 に代わって、一単語読み出し部 35、乱数発生部 36 が備えられている点が異なっている。この一単語読み出し部 35、および乱数発生部 36 の動作は図 13 に示した第 3 の実施形態におけるのと同じであり、単語辞書部 4 に格納されている単語は、乱数発生部 36 の発生する乱数に対応する単語番号に基づいて、一単語読み出し部 35 によって一単語ずつ読み出され、表示部 6 によって表示されることになる点だけが第 4 の実施形態と異なっている。なお乱数発生部 36 に代わってクロック部を設

け、クロック部が発生するクロック信号に従って、一単語読み出し部35が一単語ずつ読み出しを行うことも当然可能である。

【0041】

【発明の効果】以上詳細に説明したように、本発明によればユーザはディスプレイ画面上に自動的に表示される音声認識対象単語の中から自分の入力したい単語を見つけて単語の入力を行うことができ、マウスの操作が不必要となり、音声認識装置の操作性が向上する。また最近表示された単語だけを音声認識対象として用いることによ

って認識速度を向上させ、認識エラーを減少させることができ、音声認識装置の実用性の向上に寄与するところ

【図面の簡単な説明】

【図1】第1の発明の原理構成ブロック図である。

【図2】第2の発明の原理構成ブロック図である。

【図3】本発明における音声認識装置の第1の実施形態の構成ブロック図である。

【図4】第1の実施形態における全体処理のフローチャートである。

【図5】第1の実施形態における認識プロセスの処理フローチャートである。

【図6】第1の実施形態における複数単語の表示イメージの説明図である。

【図7】図3におけるスクロール制御部の構成ブロック図である。

【図8】図7における表示開始単語番号と表示終了単語番号のクロック信号による変化の説明図である。

【図9】音声認識装置の第2の実施形態の構成ブロック図である。

【図10】第2の実施形態における全体処理のフローチャートである。

【図11】第2の実施形態における一単語読み出しの表示イメージの説明図である。

【図12】図9における一単語読み出し部の構成ブロック図である。

【図13】音声認識装置の第3の実施形態の構成ブロック図である。

*

*【図14】第3の実施形態における全体処理フローチャートである。

【図15】図13における一単語読み出し部の構成ブロック図である。

【図16】音声認識装置の第4の実施形態の構成ブロック図である。

【図17】第4の実施形態における全体処理フローチャートである。

【図18】第4の実施形態における認識プロセスの処理フローチャートである。

【図19】図16における表示単語抽出部の構成ブロック図である。

【図20】図19におけるリングバッファの処理内容の説明図である。

【図21】音声認識装置の第5の実施形態の構成ブロック図である。

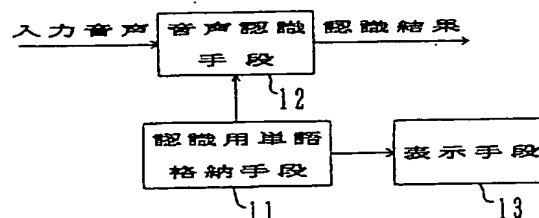
【図22】音声認識装置の従来例の構成ブロック図である。

【符号の説明】

- 1 マイク
- 2 分析部
- 3 特徴抽出部
- 4 単語辞書部
- 5 単語照合部
- 6 表示部
- 11 認識用単語格納手段
- 12, 17 音声認識手段
- 13, 18 表示手段
- 15 単語格納手段
- 16 認識用単語格納手段
- 19 表示単語抽出手段
- 21 スクロール制御部
- 22 クロック部
- 31, 35 一単語読み出し部
- 36 乱数発生部
- 41 認識用一時単語辞書部
- 42 表示単語抽出部

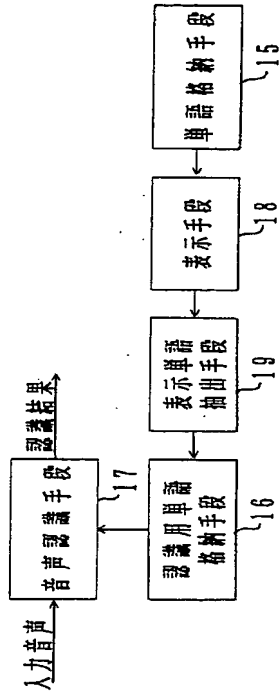
【図1】

第1の発明の原理構成ブロック図



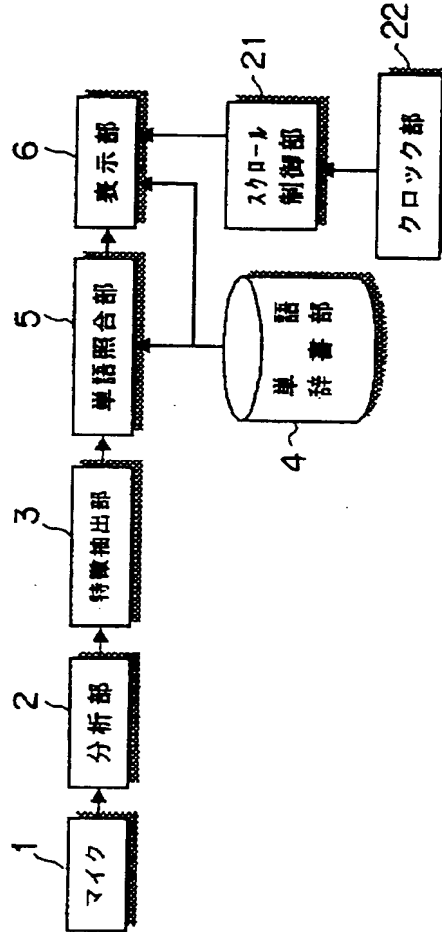
【図2】

第2の発明の原理構成ブロック図



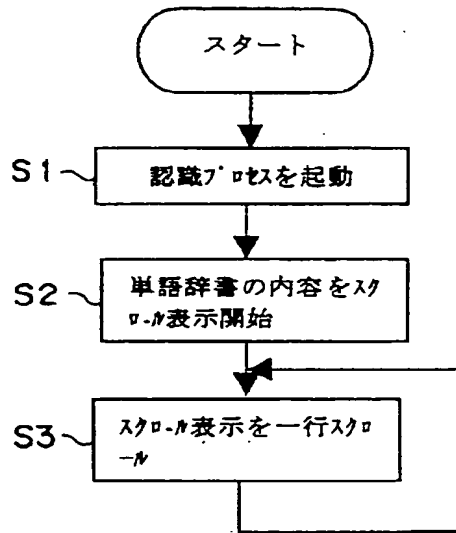
【図3】

本発明における音声認識装置の
第1の実施形態の構成ブロック図



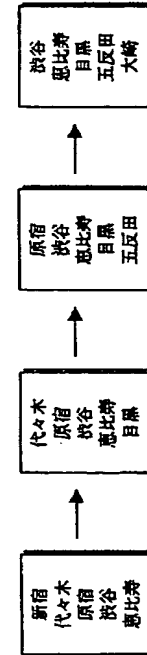
【図4】

第1の実施形態における
全体処理のフローチャート



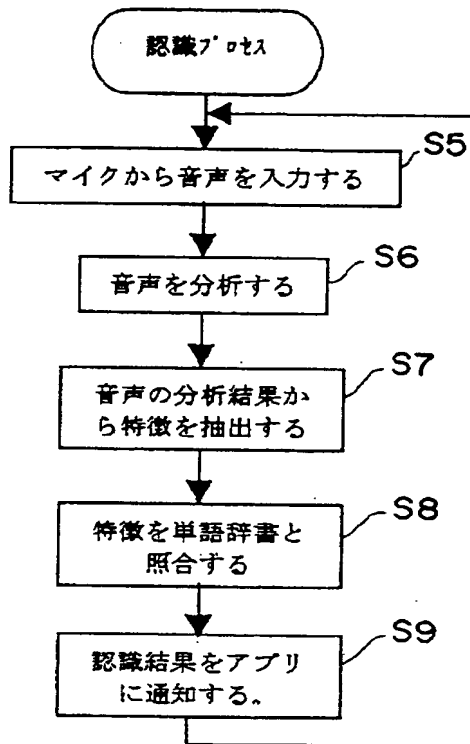
【図6】

第1の実施形態における
複数単語の表示イメージの説明図



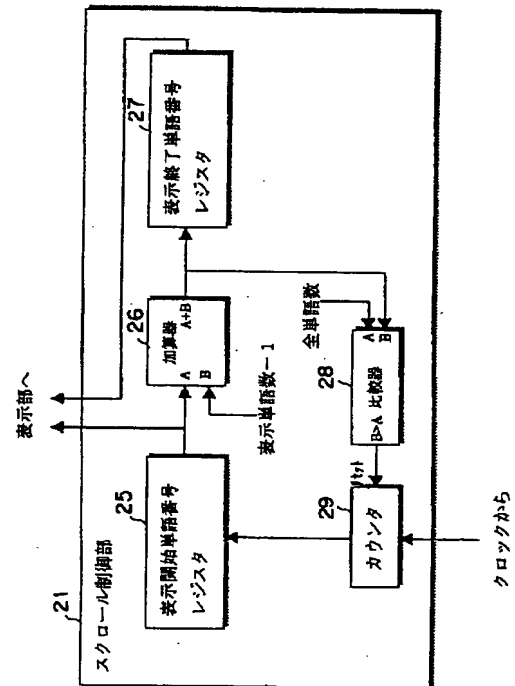
【図5】

第1の実施形態における
認識プロセスの処理フローチャート



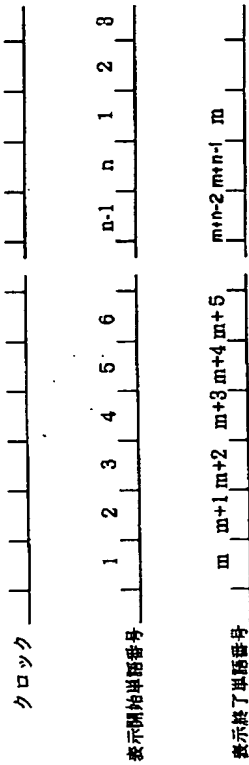
【図7】

図3における
スクロール制御部の構成ブロック図



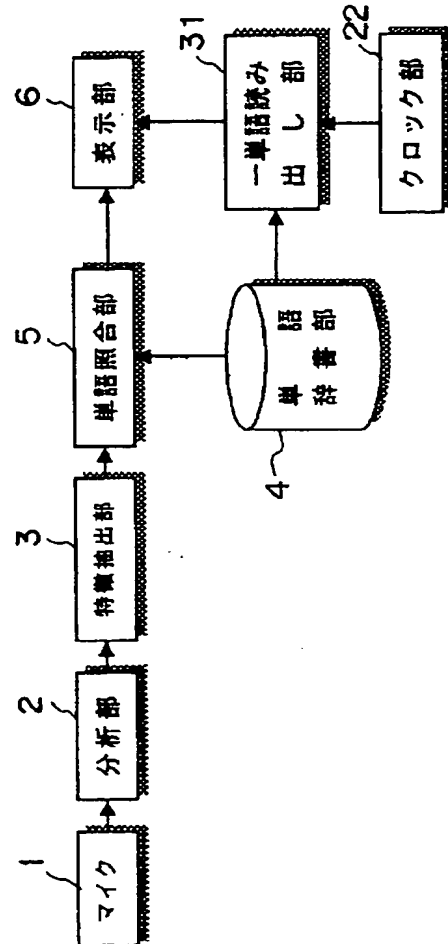
【図8】

図7における表示開始単語番号と
表示終了単語番号のクロック信号による変化の説明図



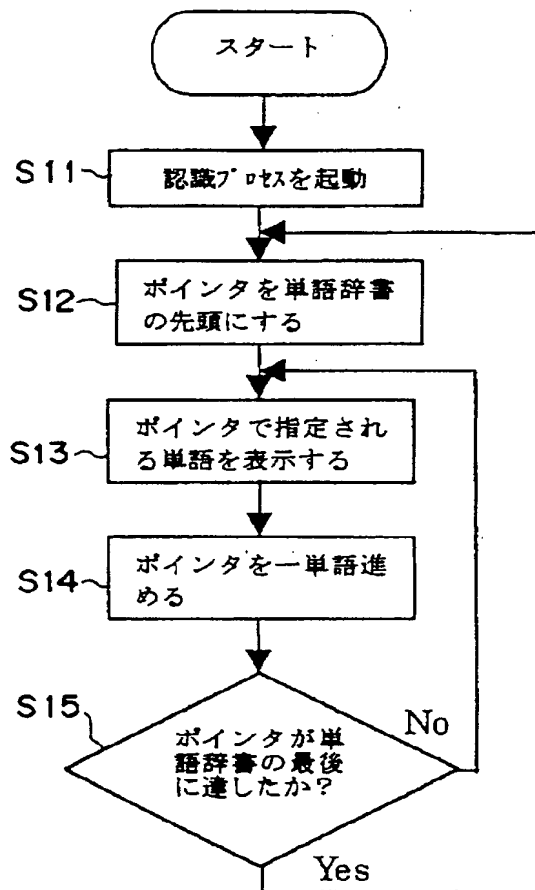
【図9】

音声認識装置の 第2の実施形態のブロック図



【図10】

第2の実施形態における
全体処理のフローチャート



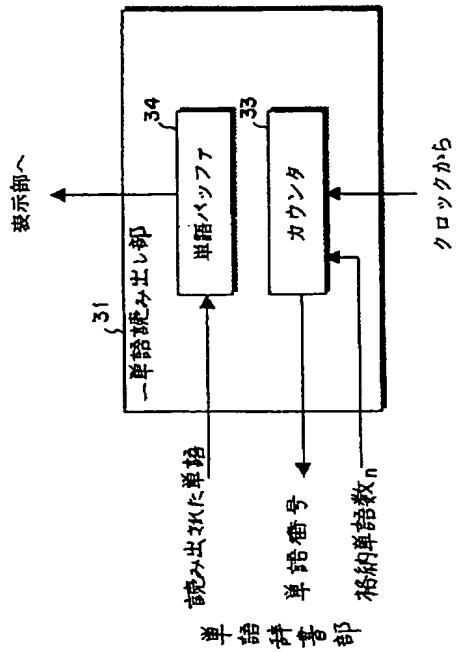
【図11】

第2の実施形態における
一単語読み出しの表示イメージの説明図



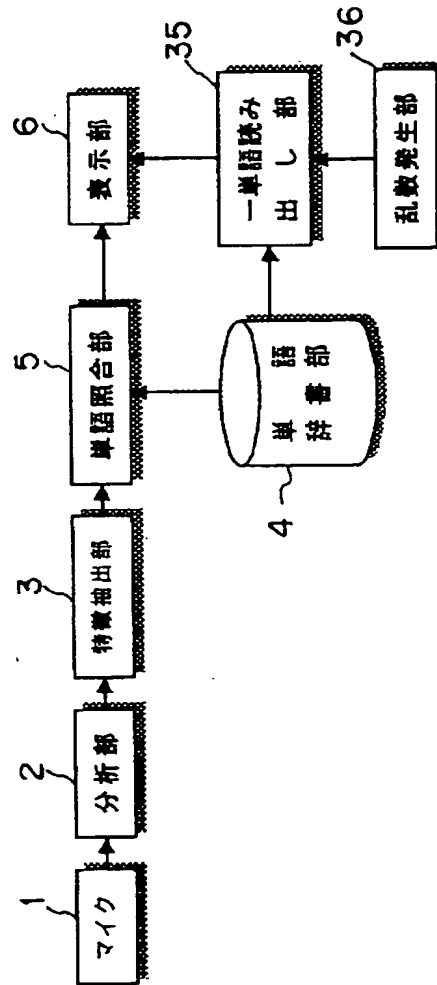
【図12】

図9における
一単語読み出し部の構成ブロック図



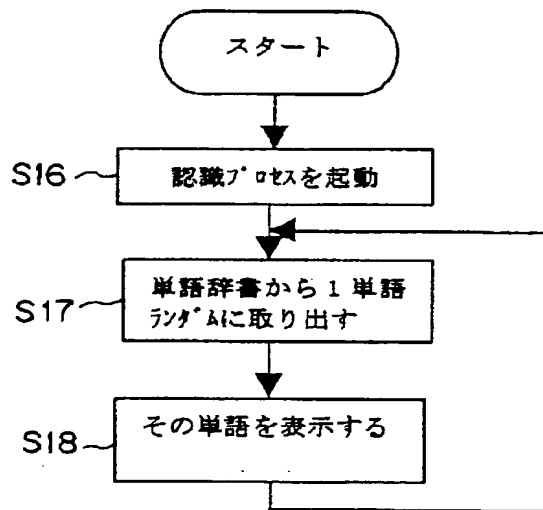
【図13】

音声認識装置の
第3の実施形態の構成ブロック図



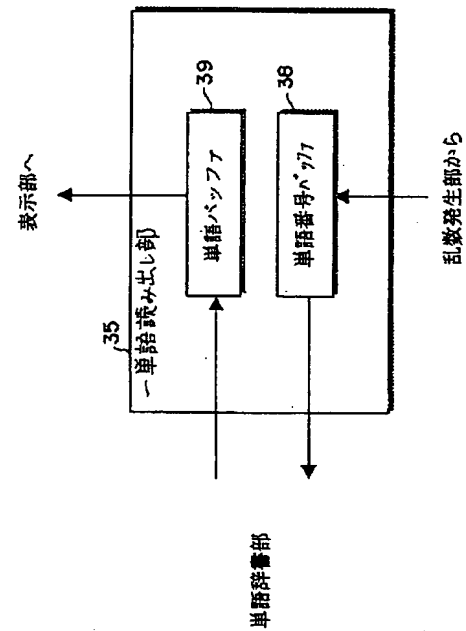
【図14】

第3の実施形態における 全体処理フローチャート



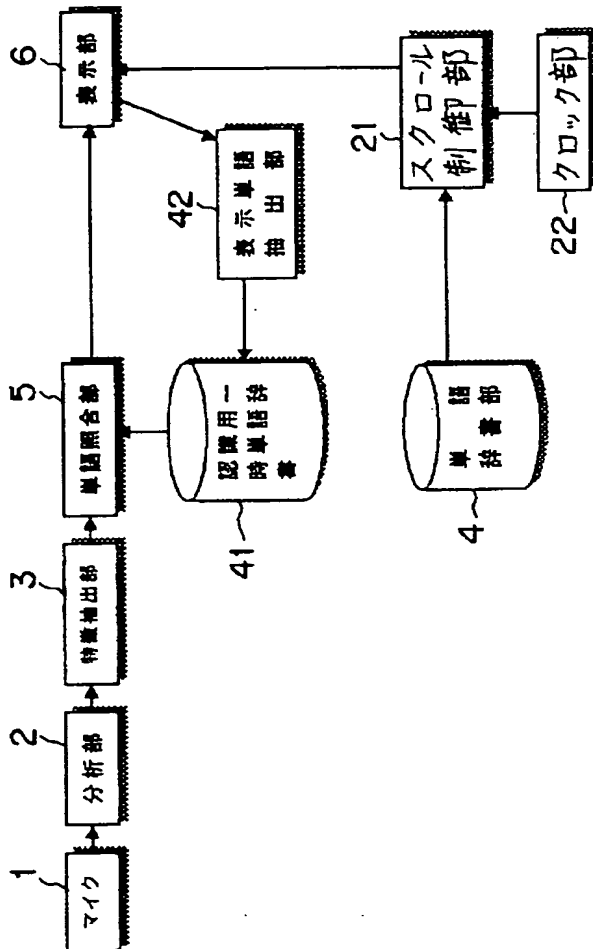
【図15】

図13における
一単語読み出し部の構成ブロック図



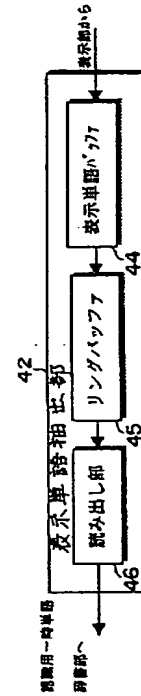
【図16】

音声認識装置の
第4の実施形態の構成ブロック図



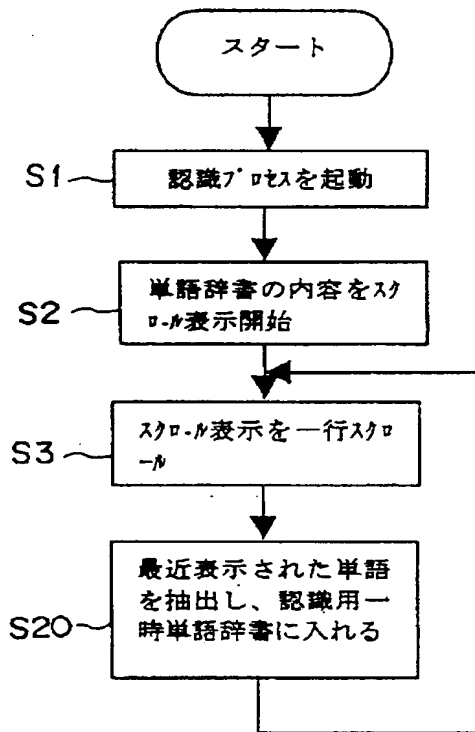
【図19】

図16における
表示単語抽出部の構成ブロック図



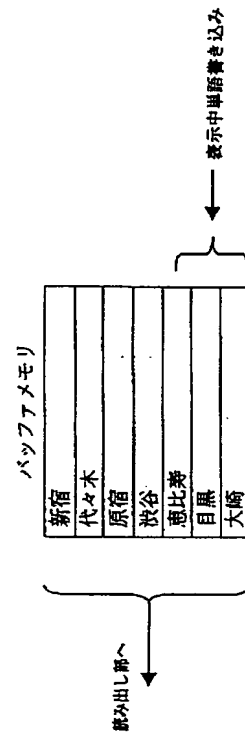
【図17】

第4の実施形態における 全体処理フローチャート



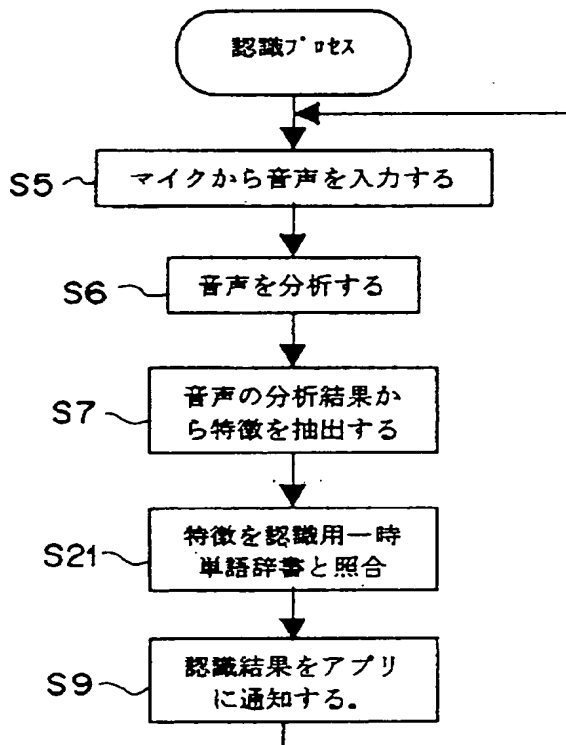
【図20】

図19における
リングバッファの処理内容の説明図



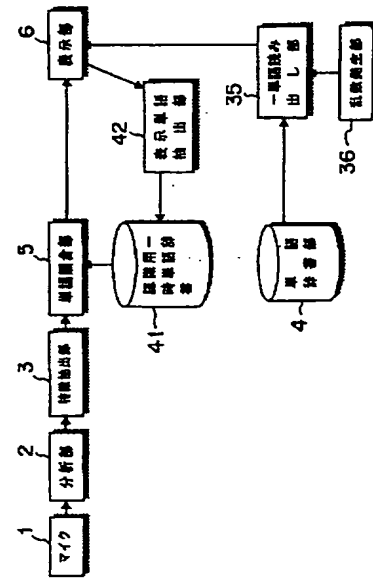
【図18】

第4の実施形態における
認識プロセスの処理フローチャート



【図21】

音声認識装置の
第5の実施形態の構成ブロック図



【図22】

音声認識装置の従来例の構成ブロック図

